ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Number Theory 118 (2006) 123–144

JOURNAL OF Number Theory

www.elsevier.com/locate/jnt

On the class groups of cyclotomic extensions in presence of a solution to Catalan's equation

Preda Mihăilescu

Mathematisches Institut der Universität Göttingen, Germany

Received 21 October 2001; revised 5 April 2005

Available online 20 December 2005

Communicated by A. Granville

Abstract

Catalan's conjecture states that the equation $x^p - y^q = 1$ has no other integer solutions but $3^2 - 2^3 = 1$. We investigate the consequences of existence of further solutions (with odd prime exponents p, q) upon the *relative* class group of the pth cyclotomic extension. We thus obtain several new results which merge into the condition

$$q \not\equiv 1 \mod p \quad \text{and} \quad p \not\equiv 1 \mod q.$$

This condition is used in the proof of Catalan's conjecture.

Il naviguait en père peinard
Sur la grand-mare des canards
Et s'app'lait « *les Copains d'abord* »,
« *les Copains d'abord* ».[1]

To Yuri, Yan and Guillaume

E-mail addresses: preda@uni-math.gdwg.de, preda@uni-paderborn.de.

[1] « Les copains d'abord », by Georges Brassens.

0022-314X/$ – see front matter
doi:10.1016/j.jnt.2005.08.011

1. Introduction

The Catalan conjecture states that the equation $X^U - Y^V = 1$ has no other solution in integer numbers except $3^2 - 2^3 = 1$; it reduces, due to results of V. Lebesgue, 1850, [Lb] and Ko Chao, 1960, [K] to the statement that

$$x^p - y^q = 1 \quad \text{with } p, q \geqslant 3 \text{ distinct primes} \tag{1}$$

has no solutions. This has been recently proved [Mih2] using methods from the theory of real cyclotomic fields.

The results which we present in this paper predate this proof and use methods which focus on the relative part of the pth cyclotomic extension $\mathbb{K} = \mathbb{Q}(\zeta_p)$. Despite the fast spreading of the final proof, we consider that these results are interesting in se in more then one respect. First, they extend upon the results of Bugeaud and Hanrot from their important paper [BH]; while doing so, they reach some natural limitations of the relative class group approach. The understanding of these limitations itself leads naturally to the solution in the real field. Briefly, the results presented here are a *missing link* between earlier research and the final proof. Furthermore, they provide a simple algebraic proof of a condition used in the final proof, and which was, prior to this paper, deduced by a long thread of deep results from Diophantine approximation, based on Baker's theory of linear forms in logarithms.

We make no longer introduction on Catalan's conjecture in this paper, and will resume to the review of the main results which we directly need for our purposes. The interested reader may consider [Ri,BH,Mi,Mih2] for in depth historical details on the research concerned with the proof of Catalan's conjecture.

We shall from now on assume that $(x, y; p, q)$ are integers, respectively primes satisfying (1). Following Cassels [Ca], we allow negative values for x and y, thus obtaining the symmetry: if $(x, y; p, q)$ is a solution of (1), then so is $(-y, -x, q, p)$. Writing h_n^- for the *relative part* of the class number of the nth cyclotomic extension, the main results of the paper are the following:

Theorem 1. *If* (1) *has a solution with odd prime exponents then* $q|h_p^-$ *and* $p|h_q^-$ *must hold.*

Theorem 2. *If* p, q *are odd primes allowing a solution of* (1)*, then*

$$\frac{q}{(p-1)^2} < 4.$$

The condition required for the proof of Catalan's conjecture, follows as a corollary of these two results:

Corollary 1. *Catalan's equation has no solution with* $\max(p, q) \equiv 1 \bmod \min(p, q)$.

The only computations required by this corollary are a table look-up for the relative class numbers of the pth cyclotomic field for $p \leqslant 137$. Finally, we prove the following independent result, which has no further applications.

Theorem 3. *If Catalan's equation has some nontrivial solution for primes $p, q > 3$, then $pq | h^+_{pq}$, where h^+_{pq} is the plus part of the class number of the pqth cyclotomic field.*

The facts which we shall further use are the following:

$$\begin{aligned} x - 1 = p^{q-1}a^q \quad &\text{and} \quad \frac{x^p - 1}{x - 1} = pv^q, \quad y = pav, \\ y + 1 = q^{p-1}b^p \quad &\text{and} \quad \frac{y^q + 1}{y + 1} = qu^p, \quad x = qbu, \end{aligned} \tag{2}$$

where a, and u, v are integers for which $(pa, v) = (qb, u) = 1$. In particular, if Catalan's equation has a solution, then

$$\frac{x^p - 1}{x - 1} = pv^q, \tag{3}$$

for some $v \in \mathbb{Z}$. This was proved by Cassels [Ca].

From this, Hyyrö [Hy] gained the following bounds:

$$\begin{aligned} |x| &\geqslant \max\bigl\{p^{q-1}(q-1)^q + 1,\ q(2p+1)\bigl(2q^{p-1}+1\bigr)\bigr\}, \\ |y| &\geqslant \max\bigl\{q^{p-1}(p+1)^p - 1,\ p(q-1)\bigl(p^{q-1}(q-1)^q + 1\bigr)\bigr\}. \end{aligned} \tag{4}$$

Finally, the Corollary 1 requires the *double Wieferich theorem* [Mih], which states that

$$\begin{aligned} p^{q-1} &\equiv 1 \mod q^2 \quad \text{and} \quad q^2 \mid x, \\ q^{p-1} &\equiv 1 \mod p^2 \quad \text{and} \quad p^2 \mid y. \end{aligned} \tag{5}$$

In (2), from $y = pav$, $(v, p) = 1$ and $p^2 | y$ it follows that $a = p \cdot a'$ and thus

$$x - 1 = p^{q-1}a^q = p^{2q-1}a'^q =: mp^{2q-1}. \tag{6}$$

2. Prerequisites and notations

The nth cyclotomic extension is denoted by $\mathbb{K}_n$ and its maximal real subfield is $\mathbb{K}_n^+$; $\zeta, \xi \in \mathbb{C}$ will be primitive pth and qth roots of unity, respectively; thus $\mathbb{K}_p = \mathbb{Q}(\zeta)$, etc. We let $P = \{1, 2, \ldots, p-1\}$, $Q = \{1, 2, \ldots, q-1\}$, and $\bar{P} = P \cup \{0\}$, $\bar{Q} = Q \cup \{0\}$. We let σ_c be the automorphism of $\mathbb{Q}(\zeta)$ with $\zeta \to \zeta^c$, for $c \in P$. The Galois group of $\mathbb{K}_p$ is $G = \mathrm{Gal}(\mathbb{Q}[\zeta]/\mathbb{Q}) = \{\sigma_c \colon c \in P\}$. Complex conjugation is denoted by $\jmath$, so $\bar{\alpha} = \jmath\alpha$.

In the investigation of Catalan's equation, we shall write $\alpha = \frac{x-\zeta}{1-\zeta}$ and $\alpha_c = \sigma_c(\alpha)$. Note that (3) becomes

$$\mathbf{N}(\alpha) = v^q$$

and there is an ideal $\mathfrak{A} = (\alpha, v) \subset \mathbb{Z}[\zeta]$ with $\mathfrak{A}^q = (\alpha)$. Furthermore,

$$\big(\sigma_a(\alpha), \sigma_b(\alpha)\big) = \big(\sigma_a(\mathfrak{A}), \sigma_b(\mathfrak{A})\big) = (1), \quad \text{for } a \neq b \in P. \tag{7}$$

Indeed, the ideal $(\sigma_a(\alpha), \sigma_b(\alpha))$, also contains $((1-\zeta^a)\alpha_a - (1-\zeta^b)\alpha_b) = \zeta^b - \zeta^a$ and it is thus at most divisible by the ramified prime $\wp$ lying above p. But $\alpha = 1 + \frac{x-1}{1-\zeta}$ and it is thus coprime to $\wp$.

2.1. Jacobi sums and the Stickelberger module

Let ℓ be a prime such that $p|(\ell-1)$, $\rho \in \mathbb{C}$ a primitive ℓth root of unity and $\chi : \mathbb{Z}/(\ell \cdot \mathbb{Z}) \to \langle\zeta\rangle$ a character of conductor ℓ and order p with $\chi(0) = 0$. Let $a, b \in P$ with $a + b \not\equiv 0 \bmod p$. The Gauss sum $\tau(\chi)$ and Jacobi sum $j(\chi^a, \chi^b)$ are defined as sums of characters by:

$$\tau(\chi) = \sum_{x \in \mathbb{Z}/(\ell\cdot\mathbb{Z})} \chi(x) \cdot \rho^x,$$

$$j\big(\chi^a, \chi^b\big) = -\sum_{x \in \mathbb{Z}/(\ell\cdot\mathbb{Z})} \chi^a(x) \cdot \chi^b(1-x). \tag{8}$$

The Jacobi sums are defined according to Lang [La], thus with a sign change with respect to their classical definition. They are related to Gauss sums by:

$$j\big(\chi^a, \chi^b\big) = -\frac{\tau(\chi^a) \cdot \tau(\chi^b)}{\tau(\chi^{a+b})}.$$

It is well known that the Gauss sum verifies $\tau(\chi) \cdot \overline{\tau(\chi)} = \ell$. If $\mathfrak{L} \supset (\ell)$ is a prime above ℓ in $\mathbb{Z}[\zeta]$, then the ideal $(j(\chi^a, \chi^b))$ will be expressed as the action of a group ring element $\psi(a,b) \in \mathbb{Z}[G]$ upon $\mathfrak{L}$. The theorem of Stickelberger [IR,Wa] computes this element: for a certain $\mathbb{Z}[\zeta_\ell] \supset \mathfrak{L} \supset (\ell)$,

$$\big(j\big(\chi^a, \chi^b\big)\big) = \mathfrak{L}^{\psi(a,b)} \quad \text{with } \psi(a,b) = \sum_{c \in P}\left(\left[\frac{c(a+b)}{p}\right] - \left[\frac{ca}{p}\right] - \left[\frac{cb}{p}\right]\right). \tag{9}$$

Here the brackets denote, as usual, the *integer part*: $[x] = \max\{a \in \mathbb{Z}\colon\ a \leqslant x\}$. The ideal $I \subset \mathbb{Z}[G]$ which is generated by such elements is called the *Stickelberger* ideal and will be presented subsequently in some detail.

We define by multiplicativity the set of *Jacobi integers* to be the subset $\mathbf{J} \subset \mathbb{Z}[\zeta]$ of products of Jacobi sums according to the above definition, and $\mathfrak{J} = \{(\mathbf{j})\colon\ \mathbf{j} \in \mathbf{J}\}$ the subset of principal ideals generated by Jacobi integers. Let $\mathfrak{A} \subset \mathbb{Z}[\zeta]$ be an ideal with $\mathbf{N}(\mathfrak{A}) = t$, such that t factors into powers of primes $\ell \equiv 1 \bmod p$. Then Stickelberger's theorem implies in particular that $\mathfrak{A}^\theta \in \mathfrak{J}$, $\forall \theta \in I$.

The following lemma is a special case adaptation of Proposition 1.2 [Jh], which relates $\mathfrak{J}$ to $\mathbf{J}$.

Lemma 1. *Let ι be the natural map $\iota : \mathbf{J} \to \mathfrak{J}$ given by $\mathbf{j} \mapsto (\mathbf{j})$. Then ι is injective. In particular, a principal ideal can be generated by at most one Jacobi integer and if for some $\alpha \in \mathbb{Z}[\zeta]$ with $\alpha \cdot \bar{\alpha} \in \mathbb{Z}$, the equality $(\alpha) = \mathfrak{j} \in \mathfrak{J}$ holds, then there is a unique Jacobi integer $\mathbf{a}$ with*:

$$\alpha = \pm\zeta^n \cdot \mathbf{a}, \quad n \in \mathbb{Z}. \tag{10}$$

Proof. Let α generate the principal ideal $\mathfrak{j} \in \mathfrak{J}$ and let $\mathbf{a} \in \mathbf{J}$ with $(\alpha) = (\mathbf{a})$: such a Jacobi integer exists by definition of $\mathfrak{J}$. The principal ideals being equal, there is a unit $\varepsilon \in \mathbb{Z}[\zeta]$ such that $\alpha = \varepsilon \cdot \mathbf{a}$. Furthermore, $\mathbf{a} \cdot \bar{\mathbf{a}} \in \mathbb{N}$ follows by multiplicativity from the property of Gauss sums and since $\alpha \cdot \bar{\alpha} \in \mathbb{Z}$, it follows that $\varepsilon \cdot \bar{\varepsilon} = 1$. By Kronecker's unit theorem, ε is a root of unity. This proves the identity (10).

We still have to prove that the Jacobi integer $\mathbf{a}$ is unique. Iwasawa shows in [Iw] (see also [IR, Exercise 13, p. 226]) that Jacobi integers $\mathbf{j}$ verify in general

$$\mathbf{j} \equiv 1 \mod (1-\zeta)^2\mathbb{Z}[\zeta]. \tag{11}$$

This property is useful for establishing the power n in (10). In particular, assuming there is a second Jacobi integer $\mathbf{a}'$, with $(\alpha) = (\mathbf{a}')$, we would have by (10) that $\alpha = \pm\zeta^{n'}\mathbf{a}'$ for some $n' \in \bar{P}$. By Iwasawa's relation, $\alpha \equiv s\zeta^n \equiv s'\zeta^{n'} \bmod \lambda^2$, where s, s' are the implicit signs for $\mathbf{a}, \mathbf{a}'$. In particular $s\zeta^n \equiv s \equiv s'\zeta^{n'} \equiv s' \bmod \lambda$ and thus $s = s'$. Also, $1 - \zeta^n \equiv n\lambda \equiv 1 - \zeta^{n'} \equiv n'\lambda \bmod \lambda^2$, so $n \equiv n' \bmod p$. Consequently, $\mathbf{a} = \mathbf{a}'$, which completes the proof. □

The notions of Gauß and Jacobi sums are easily generalized to composite orders and to characters of prime power conductor [La]. The Jacobi integers are generalized accordingly and their natural extension to *Jacobi fractions* is reflecting the extension of integral to fractional ideals [Jh]. However, we shall not need such generalizations below.

2.2. Stickelberger ideal

We introduce in this section some general and computable results about the Stickelberger ideal. We recommend, besides the classics by Washington [Wa] and Lang [La], the very rich and appealing treatment of the subject by Jha [Jh]. We shall write $d = (p-1)/2$ for convenience.

Let $\vartheta = \sum_{c=1}^{p-1} \{\frac{c}{p}\} \cdot \sigma_c^{-1} \in \mathbb{Q}[G]$ be the Stickelberger element. Stickelberger's theorem states precisely that $(\tau(\chi)) = (\mathfrak{L}^{\vartheta})$ as ideals, where $\mathfrak{L}$ lies above the conductor $\tau(\chi) \cdot \bar{\tau}(\chi)$ in an extension of degree p of $\mathbb{Q}(\zeta_p)$. The interpretation of the action of the fractional ϑ as an element of the group ring of an extension is known and we refer to [La,Wa] for details, which are of no relevance in our context.

The Stickelberger ideal is defined by $I = \mathbb{Z}[G] \cap \vartheta\mathbb{Z}[G] \subset \mathbb{Z}[G]$. Let $I' \subset \mathbb{Z}[G]$ be the ideal generated by elements of the form $n - \sigma_n$, $(n, p) = 1$. Then $I = \vartheta I'$ [Wa, §6.2]. The elements $\theta_n = (n - \sigma_n)\vartheta \in I$ are often called the *Fuchsian* elements. We also write $\theta_p = p \cdot \vartheta = \sum_{c=1}^{p-1} c \cdot \sigma_c^{-1}$. The differences $\psi_n = \theta_{n+1} - \theta_n, n \geqslant 2$ and $\psi_1 = \theta_2$ are the

Fueter elements. For the group ring elements defined in the previous section one finds: $\psi(a,b) = \sigma_a \psi_{b/a}$, where the fractional indexes are taken modulo p. Both θ_n, ψ_n form for $n = 1, 2, \ldots, d$ a $\mathbb{Z}$-base of the Stickelberger ideal [Jh]. We have for all n with $(n, p) = 1$:

$$\theta_n = (n - \sigma_n) \cdot \vartheta = \sum_{c=1}^{p-1} \left(\left[\frac{nc}{p} \right] - n \left[\frac{c}{p} \right] \right) \cdot \sigma_c^{-1} = \sum_{c=1}^{p-1} \left[\frac{nc}{p} \right] \cdot \sigma_c^{-1};$$

$$\psi_n = \sum_c \nu_c[n] \sigma_c^{-1} = \sum_c \left(\left[\frac{(n+1)c}{p} \right] - \left[\frac{nc}{p} \right] \right) \cdot \sigma_c^{-1},$$

$$\text{where } \nu_c[n] \geqslant 0 \text{ and } \nu_c[n] + \nu_{p-c}[n] = 1. \tag{12}$$

Note that the last relation implies $0 \leqslant \nu[n]_c \leqslant 1$ and

$$(1 + \jmath) \cdot \psi_n = \mathbf{N}_{\mathbb{Q}}^{\mathbb{Q}(\zeta)}, \tag{13}$$

as elements of the group ring. Since the Fueter elements generate I as a $\mathbb{Z}$-module, the identity implies herewith that

$$\forall \theta \in I,\ \exists \varsigma(\theta) \in \mathbb{Z}\colon \quad (1 + \jmath)\theta = \varsigma(\theta) \cdot \mathbf{N}_{\mathbb{Q}}^{\mathbb{Q}(\zeta)} \quad \text{and}$$

$$\forall \mathfrak{A} \subset \mathcal{O}\big(\mathbb{Q}(\zeta)\big), \quad \mathfrak{A}^{(1+\jmath)\theta} = \big(\mathbf{N}(\mathfrak{A})\big)^{\varsigma(\theta)}. \tag{14}$$

The above implies that the *weight*, i.e. the sum

$$w(\theta) = \sum_{c \in P} n_c, \quad \text{where } \theta = \sum_c n_c \sigma_c,$$

of a Stickelberger element is always a multiple of $d = (p-1)/2$. We shall also refer to $\varsigma(\theta)$ as the *relative weight of* θ. We note that

$$2 \cdot w(\theta) = w(\theta + \jmath\theta) = w\big(\varsigma(\theta) \cdot \mathbf{N}_{\mathbb{Q}(\zeta)/\mathbb{Q}}\big) = \varsigma(\theta) \cdot (p-1),$$

and thus $w(\theta) = d \cdot \varsigma(\theta)$. Note also the useful consequence of (14):

$$\Theta \in I \quad \text{and} \quad \Theta = \jmath\Theta \Rightarrow \Theta \in (\mathbf{N}_{\mathbb{K}/\mathbb{Q}}). \tag{15}$$

Indeed, if $\Theta = \jmath\Theta$, then $2\Theta = (1 + \jmath)\Theta = \varsigma(\Theta) \cdot \mathbf{N}$, so $\Theta \in \mathbb{Z}[G] \cap \frac{1}{2}(\mathbf{N}) = (\mathbf{N})$, where $(\mathbf{N}) = \mathbf{N}_{\mathbb{K}/\mathbb{Q}} \cdot \mathbb{Z}[G]$ is an ideal of the group ring. The Stickelberger ideal acts also on roots of unity and there is an additive map $\varphi : I \to \bar{P}$ such that

$$\zeta_p^{\varphi(\Theta)} = \Theta(\zeta_p), \quad \forall \Theta \in I. \tag{16}$$

In fact the map φ is determined by $\varphi(\theta_n) = \frac{n^p - n}{p} \in \bar{P}$, which follows from the Voronoi identities [IR]; we do not need this fact here.

The Stickelberger elements we consider are all *positive*, i.e. we consider the subset of I given by:

$$I^+ = \left\{ \theta = \sum_{c \in P} n_c \sigma_c \colon\ n_c \geqslant 0,\ \forall c \in P \right\} \subset I;$$

thus, for these elements we have $\varsigma(\theta) \geqslant 0$. Note in particular that all elements which arise as lifts from $\mathbb{Z}/(q \cdot \mathbb{Z})[G]/(\mathbf{N}(\mathbb{K}/\mathbb{Q}))$ to $\mathbb{Z}[G]$ are in particular positive. It will be practical to order I according to the weights and we thus define:

$$I(n) = \left\{ \theta \in I \colon\ \varsigma(\theta) = n \right\}, \quad n \geqslant 1,$$
$$\bar{I}(n) = \bigcup_{m=1}^{n} I(m). \tag{17}$$

3. The minus part

Bugeaud and Hanrot used in their important paper [BH] the archimedean version of a method due to Bilu (see also [BiH]) for their proof. Their result is $q | h_p^-$ if $q > \frac{1}{2} \cdot (1 + 1/\log(q)) \cdot p$. If $\mathbb{K}$ is the pth cyclotomic field, the number $h_p^- = h(\mathbb{K})/h(\mathbb{K}^+)$ is the *relative class number*, defined as the quotient of the class numbers of $\mathbb{K}$ and its maximal real subfield, respectively, see [Wa, Theorem 4.10]. For odd primes p, the natural map of class groups $\mathcal{C}(\mathbb{K}^+) \to \mathcal{C}(\mathbb{K})$ is injective (e.g. [Wa]) and the *relative class group* is in this case equal to the quotient group

$$\mathcal{C}_p^- = \mathcal{C}(\mathbb{K})/\mathcal{C}\big(\mathbb{K}^+\big).$$

This is the case we are interested in.[2]

We shall first adapt the method of Bugeaud and Hanrot locally and prove that the result holds unconditionally. This will yield a practical lower bound for p, q. We then generalize the archimedean approach of these authors to algebraic numbers generated by the action of the Stickelberger ideal and prove a more general result on the negative part of the class group which will imply our main result: $q \not\equiv 1 \bmod p$. The following lemma can be used in both contexts:

Lemma 2. *Let $\Theta \in \mathbb{Z}[G]$ annihilate $\mathfrak{A}$ in the relative class group $\mathcal{C}^-$. Then there is a $\nu = \nu[\Theta] \in \mathbb{Q}(\zeta)$ such that*

$$\alpha^{(1-J)\Theta} = \nu^q. \tag{18}$$

[2] In cyclotomic extensions of order divisible by more primes, there is a capitulation of order 2 and the above embedding is not injective.

Proof. By hypothesis, there are an ideal $\mathfrak{B}^+ \subset \mathbb{Q}(\zeta)^+$, so $\mathfrak{B} = \bar{\mathfrak{B}}$, and a principal ideal (ρ) with $\rho \in \mathbb{Q}(\zeta)$ such that $\mathfrak{A}^{\Theta} = (\rho) \cdot \mathfrak{B}$. Dividing by the complex conjugate of the previous equation and raising to the qth power—under consideration of $\mathfrak{A}^q = (\alpha)$—one finds the following identity between principal ideals:

$$(\alpha)^{(1-\jmath)\Theta} = \left((\rho/\bar{\rho})^q\right)$$

and $\alpha^{(1-\jmath)\Theta} = \upsilon \cdot (\rho/\bar{\rho})^q$ for some unit υ which verifies $\upsilon \cdot \bar{\upsilon} = 1$. This relation holding true under the action of the Galois group G, it follows by Kronecker's unit theorem that υ is a root of unity. The roots of unity in $\mathbb{K}_p$ are all qth powers, since $(q, 2p) = 1$, and it follows that $\alpha^{(1-\jmath)\Theta} = \nu^q$, for some $\nu = \pm\zeta^a \cdot (\rho/\bar{\rho}) \in \mathbb{Q}(\zeta)$, $a \in \mathbb{Z}$. □

We now define the *Bilu element*, which will play an important role in our subsequent investigations.

Lemma 3. *Let $\Theta \in \mathbb{Z}[G]$ be like in Lemma 2 and $\nu = \nu(\Theta)$ be such that (18) holds. Let $\omega_1, \omega_2 \in \bar{\mathbb{Z}}$ verify $\omega_1^q = \alpha^{\Theta}$ and $\omega_1/\omega_2 = \nu$, thus $\omega_2^q = \bar{\alpha}^{\Theta}$. With $\varepsilon \in \mathbb{Z}[\zeta]^{\times}$, we define $\phi = \phi(\Theta, \varepsilon) \in \mathbb{Q}(\zeta, \omega_1)$ by*

$$\phi(\Theta, \varepsilon) = \bar{\alpha}^{\Theta} \cdot (\nu - \varepsilon)^q = (\omega_1 - \varepsilon \cdot \omega_2)^q. \tag{19}$$

Then $\phi \in \mathbb{Z}[\zeta]$. Furthermore,

$$(\phi) \mid \left(\alpha^{\Theta} - \varepsilon^q \bar{\alpha}^{\Theta}\right)^q. \tag{20}$$

In particular, if $r \in P$ such that $r \cdot q \equiv 1 \bmod p$, then $\phi(1, (-\bar{\zeta}^r))$ is a unit.

Proof. We first verify that the two expressions for ϕ are equal. Indeed the first results from the second by extracting the factor ω_2^q out of the parentheses. Since the second expression is obviously in $\bar{\mathbb{Z}}$ while the first is in $\mathbb{Q}(\zeta)$, it follows that $\phi(\Theta, \varepsilon) \in \bar{\mathbb{Z}} \cap \mathbb{Q}(\zeta) = \mathbb{Z}[\zeta]$.

Let $\xi \in \mathbb{C}$ be a qth root of unity. One immediately sees that ϕ is the first factor in the product $\prod_{j=0}^{q-1}(\omega_1 - \xi^j \varepsilon \omega_2)^q = (\alpha^{\Theta} - \varepsilon^q \alpha^{\jmath\Theta})^q$. This proves (20). By applying this relation to $\phi = \phi(1, (-\bar{\zeta}^r))$ we find

$$(\phi) \mid \left(\frac{x - \zeta}{1 - \zeta} - \frac{(-\bar{\zeta}) \cdot (x - \bar{\zeta})}{1 - \bar{\zeta}}\right)^q = \left(\frac{\bar{\zeta} - \zeta}{1 - \zeta}\right)^q.$$

Since $\phi \in \mathbb{Z}[\zeta]$ and $\frac{\bar{\zeta} - \zeta}{1 - \zeta}$ is a unit, ϕ must be a unit too. □

Note that $\phi(\Theta, \varepsilon)$ is defined for annihilators Θ of $\mathfrak{A}^{1-\jmath}$. Hence, the element $\phi(1, (-\bar{\zeta}^r))$ is only defined if $\mathfrak{A}^{1-\jmath}$ is principal. It is in particular defined if $q \nmid h_p^-$.

3.1. Local approach

I owe some refinements of the current presentation of the main theorem of this section to stimulating discussions with René Schoof.

In this section we let $r \in P$ be the unique element such that $r \cdot q \equiv 1 \bmod p$ and $\mu = \alpha - 1 = \frac{x-1}{1-\zeta_p} = o(p^{2(q-1)})$, p-adically, by (6). We shall first note some properties of local pth cyclotomic extensions in relation to solutions of Catalan's equation:

Lemma 4. *Let $\mathbf{R}_p = \mathbb{Z}_p[\zeta_p]$ be the pth cyclotomic extension of $\mathbb{Z}_p$. If $\mathbf{R}_p$ contains qth roots of unity $\xi \neq 1$, then $p \equiv 1 \bmod q$ and $\xi \in \mathbb{Z}_p$.*

If ι is the natural embedding of $\mathbb{Z}[\zeta]$ in $\mathbb{Z}_p[\zeta_p]$ and $\mu_p = \iota(\mu)$, then the equation $X^q = \iota(\alpha)$ has the solutions

$$\gamma_p = \xi_0 \cdot \left(1 + \sum_{k=1}^{\infty} \binom{1/q}{k} \mu_p^k\right) \in \mathbb{Z}_p[\zeta_p], \quad \text{with } \xi_0 \in \mathbb{Z}_p,\ \xi_0^q = 1. \tag{21}$$

If $\phi = \phi(1, -\zeta^{-1/q})$ is defined, then by (21)

$$\iota(\phi) = \iota(\bar{\alpha}) \cdot \left(\frac{\gamma_p}{\bar{\gamma}_p} + \zeta^{-r}\right)^q = \left(\gamma_p + \zeta^{-r}\bar{\gamma}_p\right)^q. \tag{22}$$

Proof. The cyclotomic field $\mathbb{Q}_p(\zeta_p)$ is totally ramified, while the qth cyclotomic extension of $\mathbb{Q}_p$ is unramified [Go]; thus if $\mathbb{Z}_p[\zeta_p]$ were to contain a qth root of unity, then this should lie in $\mathbb{Z}_p$. In such case $q \mid (p-1)$, as claimed, and ξ is Galois invariant.

We now consider the equation $X^q = \iota(\alpha)$. By definition of μ_p,

$$\iota(\alpha) = \alpha_p = \frac{x - 1 + 1 - \zeta_p}{1 - \zeta_p} = 1 + \mu_p.$$

Since $v_p(\mu_p) \geqslant 1$, the Abel series in (21) converges and its sum verifies $\gamma_p^q = \alpha_p$. Furthermore, $\gamma_p/\bar{\gamma}_p$ is, up to qth roots of unity, the solution to $Y^q = \iota(\alpha^{1-J})$ and so is $\iota(\nu)$—the two numbers differ thus at most by a qth root of unity in $\mathbb{Z}_p[\zeta_p]$. Since $\mathbb{Z}_p[\zeta_p]$ is a ramified extension, such a root of unity can only exist in $\mathbb{Z}_p$. Let thus $\xi \in \mathbb{Z}_p$ be a qth root of unity, such that $\gamma_p/\bar{\gamma}_p = \xi \cdot \iota(\nu)$. The roots of unity in $\mathbb{Z}_p$ have order dividing $p-1$; thus if $p \not\equiv 1 \bmod q$, then $\xi = 1$. If $p \equiv 1 \bmod q$ let us assume that $\xi \neq 1$; since $\xi \in \mathbb{Z}_p$, this root is invariant under $\zeta_p \mapsto \zeta_p^{-1}$ (we denote this automorphism of $\mathbb{Z}_p[\zeta_p]$ also by $X \mapsto \bar{X}$). By the definition of ν at the end of the proof of Lemma 2, we have $\nu \cdot \bar{\nu} = 1$ and therefor:

$$1 = \iota(\nu) \cdot \iota(\bar{\nu}) = \left(\xi \cdot \frac{\gamma_p}{\bar{\gamma}_p}\right) \cdot \left(\xi \cdot \frac{\bar{\gamma}_p}{\gamma_p}\right) = \xi^2.$$

Since $\xi^2 = \xi^q = 1$, it follows that $\xi = 1$ in this case too and $\iota(\nu) = \gamma_p/\bar{\gamma}_p$. The definition $\phi(1, -\zeta^{-1/q}) = \alpha(\nu + \zeta^{-r})^q$ then proves the required expression for $\iota(\phi)$. □

With this we can now prove Theorem 1, the main result of this section.

Proof. We assume that the statement is false and $q \nmid h_p^-$. Then by Lemmata 2 and 3, we can define the Bilu element $\phi = \phi(1, (-\zeta^{-r}))$, which is a unit. The embedding $\iota(\phi)$ is thus given by (22), and it must also be a unit in $\mathbb{Z}_p[\zeta_p]$.

Let $r \in \mathbb{Z}$ be such that $r \cdot q \equiv 1 \bmod p^{2q-1}$. We shall show that $\iota(\phi)$ can in fact not have the norm 1, thus achieving a contradiction to the assumption $q \nmid h_p^-$.

What follows are technical computations based upon this idea.

$$\begin{aligned}\iota(\phi) &= \left(\left(1+\zeta_p^{-r}\right) + \frac{\mu_p + \zeta_p^{-r}\bar{\mu}_p}{q} + O\left(\mu_p^2\right)\right)^q \\ &= \left(1+\zeta_p^{-r}\right)^q \cdot \left(1 + \frac{\mu_p + \zeta_p^{-r}\bar{\mu}_p}{q\cdot(1+\zeta_p^{-r})} + O\left(\mu_p^2\right)\right)^q. \end{aligned} \tag{23}$$

Note that $\mu_p + \zeta_p^{-r}\bar{\mu}_p = 0$ if $r \equiv 1 \bmod p$, so we shall treat this case separately and assume first that $r \not\equiv 1 \bmod p$. Since the first factor in the above formula is obviously a unit, we must verify if the norm of the second is also 1. We develop the qth power and compute the first order term of this norm, which must vanish mod μ_p^2.

$$A = \sum_{c=1}^{p-1} \sigma_c\left(\frac{\mu_p + \zeta_p^{-r}\bar{\mu}_p}{1+\zeta_p^{-r}}\right) = (x-1)\cdot \sum_{c=1}^{p-1} \sigma_c\left(\frac{1-\zeta_p^{1-r}}{(1-\zeta_p)\cdot(1+\zeta_p^{-r})}\right), \quad \text{so}$$

$$A \equiv 0 \mod \mu_p^2. \tag{24}$$

Let $\pi = 1-\zeta_p$ be the universal uniformizer of $\mathbf{R}_p$; the running term in the sum which is the cofactor of $(x-1)$ above becomes:

$$\frac{1-\zeta_p^{1-r}}{(1-\zeta_p)\cdot(1+\zeta_p^{-r})} = \frac{1-(1-\pi)^{1-r}}{\pi\cdot(1+(1-\pi)^{-r})} = \frac{(1-r)\cdot\pi + O(\pi^2)}{\pi\cdot(2+r\pi+O(\pi^2))} = -\frac{r-1}{2} + O(\pi).$$

Note that $\sum_c \sigma_c \pi^k \equiv 0 \bmod p$ for $k > 0$; thus inserting the previous local estimate in (24) we find

$$A = -(x-1)\cdot(r-1)/2 + O\big((x-1)\cdot p\big).$$

Since $A \equiv 0 \bmod \mu_p^2$, it follows that $r \equiv 1 \bmod p$, against our assumption.

Now we assume that $r \equiv 1 \bmod p$, so $\zeta^{-r} = \bar{\zeta}$; as observed before, the first order term in (23) vanishes in this case, so we must develop the norm up to the second order. Leaving some details to the reader, the second term is in this case:

$$B = -\sum_{c=1}^{p-1} \sigma_c\left(\frac{q-1}{2q}\,\frac{\mu_p^2 + \bar{\zeta}_p\bar{\mu}_p^2}{1+\bar{\zeta}_p}\right)$$

$$= \frac{(q-1)(x-1)^2}{2q} \cdot \sum_{c=1}^{p-1} \sigma_c \left(\frac{1+\bar{\zeta}(1-\zeta)^2/(1-\bar{\zeta})^2}{(1-\zeta)^2(1+\bar{\zeta})} \right)$$

$$= \frac{(q-1)(x-1)^2}{2q} \cdot \sum_{c=1}^{p-1} \sigma_c \frac{1+\zeta}{(1-\zeta)^2(1+\bar{\zeta})} \equiv 0 \mod \mu_p^3. \tag{25}$$

Now $(1+\zeta)/(1+\bar{\zeta}) = \zeta$, so the running term in the sum (25) is $\zeta/(1-\zeta)^2 = 1/(1-\zeta)^2 - 1/(1-\zeta) = 1/\pi^2 - 1/\pi$. Given (2), (6) and the definition of μ_p, one easily sees that $v_p(\mu_p^3) - v_p((q-1)\cdot(x-1)^2) > 1$. Equation (25) implies herewith

$$\sum_{c=1}^{p-1} \sigma_c \big(1/\pi^2 - 1/\pi\big) \equiv 0 \mod p. \tag{26}$$

Note that $1/(1-\zeta)$ is a zero of the polynomial $f(X) = (X-1)^p - X^p$; thus $s_1 = \sum_c 1/\pi^{\sigma_c} = (p-1)/2$ while $s_2 = \sum_c 1/\pi^{2\sigma_c} = s_1^2 - 2\binom{p}{3}/p = -(p-1)(p-5)/12$. Finally

$$\sum_{c\in P} \sigma_c \big(1/\pi^2 - 1/\pi\big) = -\frac{p-1}{12} \cdot (6+p-5) = -\frac{p^2-1}{12}.$$

For $p \geqslant 3$ the congruence (26) is certainly impossible and a fortiori ϕ cannot be a unit. This completes the proof of the theorem. □

Corollary 2. *Catalan's equation has no solution for* $p, q < 47$.

Proof. From a table of the values of h_p^- for small primes p (e.g. [Wa]), one reads that $(p,q) = (47, 139)$ is the first (surprisingly small) pair of primes for which $q | h_p^-$ and $p \mid h_q^-$. □

Note that Eq. (3) used in Theorem 1 is $(x^p - 1)/(x-1) = py^q$ and is related to the Diophantine equation

$$\frac{x^n - 1}{x-1} = y^q \tag{27}$$

of Ljunggren and Nagell [Ri,BHM]. The methods used here can be adapted to this equation leading to results which will be gathered in a subsequent paper.

3.2. Global approach

Our next results use global bounds. We shall keep assuming that (x, y, p, q) stem from a solution to (1) and $q > p$. Corollary 2 allows the additional assumption $p \geqslant 47$.

We apply the facts exposed in the previous section on the Stickelberger ideal to the ideal $\mathfrak{A}$ generated by the presumed solution to (1). Since p is fixed, we shall from now on write I for the Stickelberger ideal $I \subset Z[\mathrm{Gal}(\mathbb{K}_p/\mathbb{Q})]$, and $I^+ \subset$ are the positive elements of I. The main consequences are:

Lemma 5. *Under the above premises,*

(i) *For every $\Theta \in I^+$ there is a $\beta_0[\Theta] \in \mathbb{Z}[\zeta]$, such that*[3] $\alpha^\Theta = \beta_0[\Theta]^q$. *The map $\Theta \mapsto \beta_0[\Theta]$ is a homomorphism.*
(ii) *Let $\Theta \in I^+$ and $\beta[\Theta] = \zeta^{\varphi(\Theta)/2q} \cdot \beta_0[\Theta]$ (with φ defined by* (16)). *Then the map β verifies*:

$$\big(\beta[\Theta]/\bar{\beta}[\Theta]\big)^q = (-1)^\Theta \cdot \left(\frac{x-\zeta}{x-\bar{\zeta}}\right)^\Theta = (-1)^\Theta \cdot \left(\frac{1-\zeta/x}{1-\bar{\zeta}/x}\right)^\Theta. \tag{28}$$

Proof. Let $\ell | v$ be a prime and $(\ell, \alpha) = \mathfrak{L} \mid (\alpha)$ be the prime above it and dividing α; this prime is unique by (7), relation which also implies that $\mathfrak{L}$ splits completely in $\mathbb{K}$. Thus $\ell \equiv 1 \bmod p$ and the factors $\mathfrak{L}^\Theta$ of (α^Θ) are generated by Jacobi integers in the restricted sense defined above. Since $\alpha \equiv 1 \bmod p$, it follows from the Iwasawa norming condition (11) that α^Θ is a Jacobi integer. The existence of β_0 follows now from Lemma 1. Indeed, it suffices to take β_0 as the unique Jacobi integer generating the principal ideal $\mathfrak{A}^\Theta$.

One easily checks that, for $n \in \mathbb{Z}$ and $\Theta, \Theta_1, \Theta_2 \in I^+$ the linearity axioms are fulfilled: $\beta_0[n \cdot \Theta] = \beta_0^n[\Theta]$ and $\beta_0[\Theta_1 + \Theta_2] = \beta_0[\Theta_1] \cdot \beta_0[\Theta_2]$. The same holds for the modified map β defined in (ii). This last map is practical for power series expansions. □

We now use the power series expansion suggested by (28) and compare it to the map β. Let $\sigma \in G, n > 0$ and consider the formal power series ring $\mathbb{Z}[\zeta, 1/q][\![T]\!]$, where G acts trivially on T. The series $f[n\sigma](T) \in \mathbb{Z}[\zeta, 1/q][\![T]\!]$ is defined by:

$$f[n\sigma] = \sum_{k \geqslant 0} \binom{n/q}{k} \sigma(-\zeta \cdot T)^k,$$

and we extend the definition to $\Theta = \sum_c n_c \sigma_c \in I$ by multiplicativity: $f[\Theta] = \prod_c f[n_c \sigma_c] \in \mathbb{Z}[\zeta][\![T]\!]$. With these definitions, the following holds:

Lemma 6. *Let $\xi \in \mathbb{C}$ be a primitive qth root of unity and $x \in \mathbb{Z}$ stem from a solution to Catalan's equation. Then*:

(i) *The explicit power series $f[\Theta](1/x)$ converges uniformly in every compact subset of $|x| > 1$ and for every $\Theta \in I$; the sum verifies*:

[3] The reader will remark that we use square brackets for maps of I which correspond to results of an action of some $\Theta \in I^+$, while we use parentheses for the values of maps from I to some numeric group.

$$f[\Theta](1/x)^q = (1-\zeta/x)^{\Theta} \quad \text{and}$$
$$\left(\frac{f[\Theta](1/x)}{f[{}_J\Theta](1/x)}\right)^q = (-1)^{\Theta}\cdot\left(\beta[\Theta]/\bar{\beta}[\Theta]\right)^q. \tag{29}$$

The map $\Theta \mapsto f[\Theta](T)$ is linear multiplicative.

(ii) *For every $\Theta \in I$ there is a unique $\kappa(\Theta) \in \mathbb{Z}/(q\cdot\mathbb{Z})$, such that*

$$(-1)^{\Theta}\cdot\left(\beta[\Theta]/\bar{\beta}[\Theta]\right) = \xi^{\kappa(\Theta)}\cdot\left(\frac{f[\Theta](1/x)}{f[{}_J\Theta](1/x)}\right). \tag{30}$$

As a map $I \to \mathbb{Z}/(q\cdot\mathbb{Z})$, κ is linear additive. Furthermore,

$$\kappa[{}_J\Theta] = -\kappa(\Theta), \quad \forall\Theta \in I. \tag{31}$$

Proof. Note that $f[n\sigma](T)$ is the formal Abel (generalized binomial) series expansion for $(1-\zeta\cdot T)^{n/q}$. It is a result from elementary analysis, that the series converges in the complex plane uniformly for $T = 1/x$ in all compact subdomains of $|x| > 1$; thus in particular for x stemming from a solution of Catalan's equation. Together with the definition (28) of β, this proves (i). We remark that $\beta[\Theta]/\bar{\beta}[\Theta] \in \mathbb{Q}(\zeta)$, while this is not necessarily true for $f[\Theta](1/x)/f[{}_J\Theta](1/x) \in \mathbb{C}$; however, since both have the same qth power and $\mathbb{C}$ is a field, they can only differ by a qth root of unity—which leads to the definition of $\kappa(\Theta)$. The additivity of κ is simply verified, while the important symmetry condition (31) is a consequence of complex conjugation being a continuous map on $\mathbb{C}$. □

The next lemma prepares some technical estimates on f and β:

Lemma 7. *Notations being like in the previous lemma, we let $\Theta = \sum_c n_c\sigma_c \in I^+$ with $0 \leqslant n_c < q$, so $0 < w(\Theta) < q(p-1)$ and write $\eta[\Theta] = \sum_c n_c\sigma_c(\zeta)$. Then*

(i) *The second order remainder of the series $f[\Theta](1/x)$ is, for $|x| > 1$:*
$R_2[\Theta](1/x) = f[\Theta](1/x) - (1-\eta[\Theta]/x)$ and it is bounded by:

$$\left|R_2[\Theta](1/x)\right| \leqslant \frac{1}{2}\left(\frac{w(\Theta)/q+2}{|x|}\right)^2. \tag{32}$$

(ii) *The remainder $|\delta(\Theta)| = \left|\frac{f[\Theta](1/x)}{f[{}_J\Theta](1/x)} - 1\right| < \frac{3w(\Theta)}{q|x|}$.*

(iii) *For $\Theta \in I^+$,*

$$\left|(-1)^{\Theta}\beta[\Theta]/\bar{\beta}[\Theta] - 1\right| < \begin{cases} \frac{3\cdot w(\theta)}{q\cdot|x|}, & \text{if } \kappa(\theta) = 0, \\ 3, & \text{otherwise.} \end{cases} \tag{33}$$

Proof. A power series $\sum_{k=0}^{\infty} a_k T^k$ with complex coefficients is *dominated*[4] by the series $\sum_{k=0}^{\infty} A_k T^k$ with nonnegative real coefficients if $|a_k| \leqslant A_k$ for $k \geqslant 0$. The relation of dominance is preserved by addition and multiplication of power series.

The binomial series $f[n\sigma] = (1-\sigma\zeta T)^{n/q} = \sum_{k=0}^{\infty} \binom{r}{k}(-\sigma\zeta)^k T^k$ with coefficients $a_k = \binom{r}{k}(-\sigma\zeta)^k$ is dominated by $(1-T)^{-n/q} = \sum_{k=0}^{\infty}(-1)^k\binom{-n/q}{k}T^k$. Indeed $A_k = (-1)^k \cdot \binom{-n/q}{k} \geqslant 0$ and $\left|\binom{n/q}{k}\right| \leqslant \left|\binom{-n/q}{k}\right|$, while $|(-\sigma\zeta)^k| = 1$. It follows by multiplicativity that $f[\Theta](T)$ is dominated by $(1-T)^{-w(\Theta)/q}$ and so are the partial sums and remainders. Next we write $w(\Theta) = q\cdot h - n$ with $0 \leqslant n < q$, so $pq \geqslant w(\Theta) + q > hq$ and $h < p$ and also $(h+1)^2 < |x|$, by (4). We estimate the binomial coefficient

$$\left|\binom{-w(\Theta)/q}{k}\right| = \left|\binom{-h+n/q}{k}\right| = \left|\frac{(-h+n/q)\cdot\cdots\cdot(-h+n/q-(k-1))}{k!}\right|$$
$$\leqslant \frac{h(h+1)\cdot\cdots\cdot(h+k-1)}{k!} = \binom{h+k-1}{k}.$$

Note also that for $k \geqslant 2$ we have

$$\binom{h+k}{k+1} = \frac{h+k}{k+1}\binom{h+k-1}{k} < (h/3+1)\cdot\binom{h+k-1}{k} < (h/3+1)^{k-2}\binom{h+1}{2}.$$

With this, and writing $m = h/3+1$, the error term estimates to:

$$\left|R_2[\Theta](1/x)\right| \leqslant \sum_{k=2}^{\infty}\binom{h+k-1}{k}\frac{1}{|x|^k} < \frac{h(h+1)}{2|x|^2}\cdot\sum_{k=0}^{\infty}\left(m/|x|\right)^k$$
$$= \frac{h(h+1)}{2|x|^2}\,\frac{|x|}{|x|-m} = \frac{h(h+1)}{2|x|(|x|-m)}.$$

But $(h+1)^2 < |x|$, so $\frac{h}{|x|-m} < \frac{h+1}{|x|}$ and the above becomes:

$$\left|R_2[\Theta](1/x)\right| < \frac{1}{2}\left(\frac{h+1}{|x|}\right)^2,$$

which is (i), since $h+1 < w(\Theta)/q+2$.

For the proof of (ii), we combine (29) to (32). Let $A = f[\Theta](1/x)$, $B = -\frac{\eta[\Theta]}{qx}$ and $R_2 = R_2[\Theta](1/x)$. Then $A = 1+B+R_2$ and, by (32),

$$|R_2 - \bar{R}_2| \leqslant |R_2| + |\bar{R}_2| = 2|R_2| < C = \left(\frac{w(\Theta)/q+2}{|x|}\right)^2,$$

[4] This elegant estimation technique has been suggested by Yuri Bilu in a similar context.

and

$$|B-\bar{B}| = \frac{1}{q|x|}\cdot\left|\sum_c n_c\left(\zeta^c-\bar{\zeta}^c\right)\right| < 2\cdot W = 2\cdot\frac{w(\Theta)}{q|x|}.$$

We have the lower bound:

$$\left|\bar{f}[\Theta](1/x)\right| = |1+\bar{B}+\bar{R}_2| \geqslant 1-W-C.$$

Assembling the recent estimates, we find:

$$\delta(\Theta) = \left|\frac{f[\Theta](1/x)}{\bar{f}[\Theta](1/x)}-1\right| = \left|\frac{B-\bar{B}+R_2-\bar{R}_2}{\bar{f}[\Theta](1/x)}\right| \leqslant \frac{2\cdot W+C}{1-W-C}.$$

Finally, $W = w(\Theta)/q|x| < (q(p-1))/(qp^{q-1}) < 1/p^{q-2} \leqslant 1/9$ certainly holds, so $|(2\cdot W+C)/(1-W-C)| < 3\cdot W$, which implies claim (ii). Indeed, we have

$$C < \frac{1}{q}\frac{w(\Theta)}{q|x|} = W/q < \frac{1}{q(q+1)};$$

the first inequality amounts, after reducing the fractions, to $1+2q/w(\Theta) < \sqrt{|x|/w(\Theta)}$, which is obvious from (4) and the bound $w(\Theta) < (p-1)q$. The second is $(q+1)\cdot w(\Theta) < (p-1)q(q+1) < |x|$. Thus

$$\frac{2\cdot W+C}{1-W-C} < \frac{(2+1/q)\cdot W}{1-W(q+1)/q} < 3\cdot W,$$

as claimed; note that the last inequality is equivalent to $2+1/q < 3(1-W(q+1)/q)$, so $3W+\frac{3W+1}{q} < 1$, which is true for $W < 1/9$, $q \geqslant 5$, as we assumed.

We can now combine (ii) with (30) and prove the fundamental bounds (iii). We have:

$$\begin{aligned}\left|(-1)^{\Theta}\beta[\Theta]/\bar{\beta}[\Theta]-1\right| &= \left|\xi^{\kappa(\Theta)}\cdot\frac{f[\Theta](1/x)}{f[J\Theta](1/x)}-1\right| = \left|\xi^{\kappa(\Theta)}-1+\delta(\Theta)\cdot\xi^{\kappa(\Theta)}\right| \\ &< \left|\xi^{\kappa(\Theta)}-1\right| + \left|\delta(\Theta)\right|.\end{aligned}$$

If $\kappa(\Theta)=0$, the bound on $\delta(\Theta)$ implies the respective claim in (33). Otherwise

$$\left|(-1)^{\Theta}\beta[\Theta]/\bar{\beta}[\Theta]-1\right| < 2+\frac{2w(\Theta)}{q|x|} < 3.$$

This completes the proof of (33). □

As mentioned previously, our main result generalizes the archimedean method of Bugeaud and Hanrot. For $\Theta \in I^+$, we use the map $\beta[\Theta]$ defined above which yields the slightly modified Bilu-element:

$$\phi[\Theta] = \beta[\Theta] - (-1)^{\Theta}\bar{\beta}[\Theta] = \bar{\beta}[\Theta]\cdot\left(\frac{\beta[\Theta]}{\bar{\beta}[\Theta]} - (-1)^{\Theta}\right)$$
$$= (-1)^{\Theta}\cdot\bar{\beta}[\Theta]\cdot\left(\xi^{\kappa(\Theta)}\cdot\frac{f[\Theta](1/x)}{f[\jmath\Theta](1/x)} - 1\right). \tag{34}$$

Note that ϕ differs from the previously defined Bilu elements in two ways: first, since $\Theta \in I$ annihilates not only $\mathfrak{A}^{1-\jmath}$ but also $\mathfrak{A}$, no qth power is required. Second, $\beta[\Theta]/\bar{\beta}[\Theta]$ is only a qth root of $\left(\frac{\alpha}{\bar{\alpha}}\right)^{\Theta}$ up to a $2p$th root of unity.

We have in (33) a very efficient tool for estimating the norm of $\phi[\Theta]$. It is evident that for doing this for a fixed $\Theta \in I^+$, the vector $\mathcal{K}[\theta] = (\kappa(\sigma\Theta))_{\sigma_{\in G}} \in (\mathbb{Z}/(q\cdot\mathbb{Z}))^{p-1}$ is very indicative. In particular the number of zero entries in this vector, which is an even number due to (31) reflects, by (33), the number of conjugates of $\phi[\Theta]$ which are small in absolute value. We shall denote by

$$2\mathfrak{z}(\theta) = \sharp\big\{\sigma \in G\colon\ \kappa(\sigma\Theta) = 0\big\},$$

the number of zero entries in the vector $\mathcal{K}(\Theta)$: $\mathfrak{z}$ counts the *pairs* of vanishing Galois exponents. The next theorem establishes a relation between the number of vanishing exponents and the weight of Θ, relation which depends on p and q:

Theorem 4. *If $\Theta \in I^+ \setminus (\mathbf{N})$ has weight $w(\Theta) < q(p-1)$ and is such that $\mathfrak{z}(\Theta) > 0$, then*

$$w(\Theta)\cdot(p-1) > \big(2\mathfrak{z}[\Theta] - 1\big)\cdot q. \tag{35}$$

Proof. We shall estimate in different ways the absolute value of the norm $N = |\mathbf{N}_{\mathbb{Q}}^{\mathbb{Q}(\zeta)}(\phi[\Theta])|$. First we show that $\phi[\Theta] \neq 0$. Indeed, if this were not the case, then

$$1 = \big((-1)^{\Theta}\cdot\beta[\Theta]/\bar{\beta}[\Theta]\big)^q = \left(\frac{x-\zeta}{x-\bar{\zeta}}\right)^{\Theta},$$

and thus

$$\big((1-\zeta)\alpha\big)^{\Theta} = \big((1-\bar{\zeta})\bar{\alpha}\big)^{\Theta}.$$

By dividing out a power of the ramified prime, this implies the equality of principal ideals: $(\alpha^{\Theta}) = (\alpha^{\jmath\Theta})$. Since $\mathbf{N}(\alpha) = v^q > 1$, α is not a unit; by (7), $(\sigma_a(\alpha), \sigma_b(\alpha)) = (1)$ implies $a = b$. The equality of principal ideals implies herewith that $\Theta = \jmath\Theta$ and Θ is a multiple of the norm as a consequence (15). This contradicts the hypothesis and it follows that $\phi[\Theta] \neq 0$ and we have the trivial lower bound $|\mathbf{N}(\phi[\Theta])| \geqslant 1$.

The norm estimates will involve the value of the pth cyclotomic polynomial $\Phi_p(x)$ and we note that

$$\big|\Phi_p(x)\big| = \big|x^{p-1} + x^{p-2} + \cdots + 1\big| < \frac{4}{3}|x|^{p-1}, \quad \text{for } |x| > 4.$$

We use now (28) and (33) and write $\nu = \beta[\Theta]/\bar{\beta}[\Theta]$ and $w = w(\Theta) < q(p-1)$; a simple computation shows that the norm

$$N = \mathbf{N}_{\mathbb{Q}}^{\mathbb{Q}(\zeta)}\big(\phi[\Theta]\big) = \Phi_p(x)^{w/q} \cdot \prod_{c \in P} \big((-1)^{\Theta}\sigma_c(\nu) - 1\big)$$

is bounded by:

$$\begin{aligned} 1 \leqslant |N| &\leqslant \left|\Phi_p^{w/q}(x)\right| \cdot 3^{p-1} \cdot \left(\frac{w}{q \cdot |x|}\right)^{2\mathfrak{z}} \\ &< \big(4/3 \cdot |x|^{p-1}\big)^{w/q} \cdot 3^{p-1} \cdot \left(\frac{p-1}{|x|}\right)^{2\mathfrak{z}}, \end{aligned}$$

since $w/q \leqslant p-1$. Let us separate the constants from powers of $|x|$ in the above inequality; we find:

$$\big(|x|\big)^{2\cdot\mathfrak{z}-(p-1)w/q} < 4^{p-1} \cdot (p-1)^{2\mathfrak{z}}. \tag{36}$$

It follows from (4) and $q > p$ that $|x| > (4\cdot(p-1))^{p-1} > 4^{p-1}\cdot(p-1)^{2\mathfrak{z}}$ and the exponent of $|x|$ in (36) must be < 1, thus $2\mathfrak{z} - 1 < (p-1)w/q$, which is the claim (35). □

The following is a simple consequence of the theorem:

Corollary 3. *If $\Theta \in I^+$ is such that $(p-1)\cdot w(\Theta) < q$ then $\mathfrak{z}(\Theta) = 0$.*

Proof. We have, by (35) and the definition of Θ, the following inequalities:

$$\big(2\mathfrak{z}(\Theta) - 1\big) \cdot q < (p-1)w(\Theta) < q,$$

which combine to $2(\mathfrak{z}(\Theta) - 1) \cdot q < 0$; since $\mathfrak{z}$ is an integer valued function, this implies that $\mathfrak{z}(\Theta) = 0$ as claimed. □

The idea of the next central proposition is that if q is too large, then there will be numerous $\Theta \in I^+$ verifying the premise of Corollary 3; namely sufficiently many in order to prove that $\mathfrak{z}$ cannot vanish for all. Let $d = (p-1)/2$, as previously, and remember that the Fueter elements form a $\mathbb{Z}$-base of the Stickelberger ideal.

Lemma 8. *Let*

$$s = \left[\frac{q}{(p-1)^2}\right].$$

Then

(i) *If $s > 0$ then $\mathfrak{z}(\Theta) = 0$ for all $\Theta \in \bar{I}(2s)$.*

(ii) *Furthermore, the restriction of the Galois exponent map* $\kappa : I \to \mathbb{Z}/(q \cdot \mathbb{Z})$ *to* $\bar{I}(s) \bmod q$ *is injective and nonvanishing. In particular* $|I(s) \bmod q| < q$.

Proof. We prove (i), which is a consequence of Corollary 3. Indeed, let $\Theta \in \bar{I}(2s)$, so by definition of $\bar{I}(2s)$, we have $w(\Theta) \leqslant 2s \cdot \frac{p-1}{2} = s(p-1)$ and thus $(p-1) \cdot w(\Theta) \leqslant s(p-1)^2$. By definition of s and since $q/(p-1)^2$ is not an integer, we also have $0 < s < q/(p-1)^2$, hence

$$(p-1)w(\Theta) \leqslant s(p-1)^2 < q,$$

and Corollary 3 implies that $\jmath(\Theta) = 0$. This means that the Galois exponent map never vanishes on $\bar{I}(2s)$. We now show that it is also injective for $\Theta \in \bar{I}(s)$. If it were not injective, then there are $\Theta_1 \neq \Theta_2 \in \bar{I}(s) \bmod q$ such that $\kappa(\Theta_1) = \kappa(\Theta_2)$. But then, taking $\Theta = \Theta_1 + \jmath \cdot \Theta_2$, we obviously have $\Theta \in \bar{I}(2s)$ while $\kappa(\Theta) = \kappa(\Theta_1) - \kappa(\Theta_2) = 0$ and thus $\jmath(\Theta) > 0$. This contradicts (i) and completes the proof of the lemma. □

The next lemma provides a crude estimate of the number of distinct elements $\Theta \in (I(m) \bmod q\mathbb{Z}[G])$, for $1 \leqslant m < q$.

Lemma 9. *Notations being like above,*

(i) *Let* $0 < m < q$ *be an integer,*

$$A(m) = \left\{ \vec{a} = (a_1, a_2, \ldots, a_d) \subset \mathbb{Z}^d_{\geqslant 0} \colon \sum_{i=1}^{d} a_i = m \right\} \quad \textit{and}$$

$$\Psi : A(m) \to I \quad \textit{such that} \quad \vec{a} \mapsto \Theta = \sum_{i=1}^{m} a_i \psi_i.$$

We let $\Psi' = \Psi \bmod q\mathbb{Z}[G]$. *Then* Ψ' *is injective.*

(ii) *For* $0 < m < q$,

$$\sharp\big(I(m) \bmod q\mathbb{Z}[G]\big) \geqslant \binom{d+m-1}{m}. \tag{37}$$

Proof. The module $I \bmod q\mathbb{Z}[G]$ may have smaller rank then d; this happens exactly when $q \mid h_p^-$ [Jh]. The idea of this lemma is thus to use the $\mathbb{Z}$-base provided by the Fueter elements in order to construct a sufficiently large family of elements in I which are mutually distinct modulo q. The resulting estimate is less then optimal, but works around the problem of estimating the $\mathbb{F}_q$-base of $I \bmod q\mathbb{Z}[G]$; the result is sufficient for our purpose.

Let $A(m)$ be defined by (i) and $\vec{a} \neq \vec{a}' \in A(m)$. Since the Fueter elements are a $\mathbb{Z}$-base of I, the images $\Theta = \Psi(\vec{a})$, $\Theta' = \Psi(\vec{a}') \in I$ are distinct. We show that they are also distinct

modulo q. Let $\vec{a} \in A(m)$ and $\Psi(\vec{a}) = \sum_{c\in P} n_c\sigma_c$. The coefficients of the Fueter elements are 0 or 1 and thus $0 \leqslant n_c \leqslant \sum_{i=1}^{d} a_i = m < q$. Let now $\vec{a}, \vec{a}' \in A(m)$ and consider

$$\Theta = \Psi(\vec{a}) - \Psi(\vec{a}') = \sum_c \big(n_c - n'_c\big)\sigma_c.$$

If $\Theta \equiv 0 \bmod q\mathbb{Z}[G]$, then $n_c \equiv n'_c \bmod q$ for all $c \in P$. But we showed that $0 \leqslant n_c, n'_c < q$, so this implies $n_c = n'_c$ and $\Psi(\vec{a}) = \Psi(\vec{a}')$. Thus $\vec{a} = \vec{a}'$ and $\Psi \bmod q$ is injective, which is claim (i). By a simple combinatorial exercise, one has $\sharp A(m) = \binom{d+m-1}{m}$ and the inequality (37) follows from the injectivity of Ψ'. □

We can now prove Theorem 2: $\frac{q}{(p-1)^2} < 4$.

Proof. Let $s = \big[\frac{q}{(p-1)^2}\big]$. Certainly, if $s \leqslant 3$, then there is nothing to prove, so we assume that $s > 3$ and note that by definition, $s < q$. Thus, the estimate (37) holds. Combining this with (ii) of Lemma 8, we must have

$$\binom{d+s-1}{s} < \big|\bar{I}(s) \bmod q\big| < q. \tag{38}$$

The definition of integer parts yields directly: $s \cdot (p-1)^2 < q < (s+1)\cdot(p-1)^2$. Thus the assumption that

$$\binom{d+s-1}{s}\Big/(s+1) = \binom{d+s-1}{s+1}\Big/(d-1) > (p-1)^2 \tag{39}$$

implies $\binom{d+s-1}{s} > (s+1)(p-1)^2 > q$, a contradiction to (38). We shall show by induction that (39) holds for $s \geqslant 4$ and $p > 45$.

Let $s = 4$. Since $\binom{d+s-1}{s} > d^s/s!$ and an easy calculation shows that

$$\left(\frac{p-1}{2}\right)^4 > 5!\cdot(p-1)^2 \quad \text{for } p \geqslant 45,$$

it follows that (39) holds for $s = 4$, $p > 45$. We have the recurrence $\binom{d+s}{s+2} = \frac{d+s}{s+2}\cdot\binom{d+s-1}{s+1}$. For $p \geqslant 5$, $\frac{d+s}{s+2} \geqslant 1$ and a fortiori, if $p > 45$, the inequality (39) holds for all $s \geqslant 4$. If $p > 45$, the assumption (39) is thus false for $s > 3$. Since we proved that in solutions of Catalan's equation $p > 45$ must hold, it follows that $s \leqslant 3$ which implies the claim. □

The Corollary 1 follows:

Proof. By Corollary 2, if there is a solution of (1) with $q > p$, then $p \geqslant 47$. The Theorem 2 then implies that $q < 4(p-1)^2$. However, if $q \equiv 1 \bmod p$, then $q \equiv 1 \bmod p^2$ by the double Wieferich condition (5). Since q is an odd prime, we should in fact have $q = 2kp^2 + 1$, for some $k > 0$; we shall show that $k \neq 1$, so $q > 4p^2$ in contradiction with Theorem 2.

Note[5] that $p = 3$ or $p^2 \equiv 1 \bmod 3$. Since q is prime, it follows that $k \not\equiv 1 \bmod 3$ except when $p = 3$, case in which $q = 19$. This case is discarded by Corollary 2, so $k > 1$. There are thus no solutions of (1) for $q \equiv 1 \bmod p$. The statement follows since we chose $q > p$ and thus $\max(p, q) = q$ and $\min(p, q) = p$. □

Remark 1. The proof of Theorem 1 is in principle a generalization of the global method of Bugeaud and Hanrot in order to take advantage of the Stickelberger annihilation. Just like for the proof [Mi] of (5), the use of Stickelberger as annihilators is incidental, being handy. It is not stringent, and more general annihilators of the negative part could be used. For (5) this has been observed by Puchta [Pu], leading however to the same result. The same is true in this case. More important, the present proof is based on quite simple estimates of both $I(m) \bmod q$ and of $\eta[\Theta]$. However, even using much stronger—and harder to prove—estimates and techniques, it seems that the best improvement one can achieve is replacing the constant 4 in Proposition 2 by the constant 2 or slightly less. It is striking that the unspecific results on linear forms in logarithms yield stronger results then this analysis which appears to take more advantage of the particular properties of Eq. (1). This certainly speaks in favor of the power of that method. Yuri Bilu gives in [Bi] an alternate proof of Theorem 2, using heights and Liouville's inequality. As a result, the lower bounds for $|x|$ used are looser then the ones provided by (4).

The methods using the negative part of the class field have natural limitations, which arise from the Galois exponents: they could only be dealt with by using combinatorial arguments, which seriously reduce the power of the method. It is thus natural to overcome this problem by moving the argument to the real subfield of $\mathbb{K}_p$, where the Galois exponents naturally vanish. This is done in the proof of Catalan's conjecture [Mih2], which uses Corollary 1.

4. Reflection

I start with a *theorem which I was not yet able to prove, hoping someone else might be luckier*[6]—it is a slight generalization of Vandiver's conjecture and the rest of this section will show that it implies Catalan.

Conjecture 1. If p, q are two distinct odd primes, then

$$pq \nmid h^+_{pq}.$$

Based on Theorem 1, we are able to apply to Catalan's equation the methods specific for the cyclotomic investigation of the Second Case of Fermat's last theorem. Both relations (2) and (5) hold symmetrically in p, q. We shall assume in what follows, that x, v stem from a nontrivial solution of (3) for odd primes p, q. The following lemma is an adaptation of Lemma 9.1 [Wa] to the present context.

[5] I owe this very useful remark to M. Mignotte.

[6] These are the words with which Catalan recommended his conjecture in his 1844 note in Crelle.

Lemma 10. *If $\alpha = \frac{x-\zeta}{1-\bar{\zeta}}$, then the extension $\mathbb{K}_{pq}((\alpha/\bar{\alpha})^{1/q})$ is unramified over $\mathbb{K}_{pq}$.*

Proof. Let ξ be a qth primitive root of unity and $\mathbb{K} = \mathbb{K}_{pq} = \mathbb{Q}(\zeta, \xi)$. Then $a = -\bar{\zeta} \cdot \alpha/\bar{\alpha} \in \mathbb{K}^*$ and the ideal $(\alpha) = \mathfrak{A}^q$ is the qth power of an ideal in $\mathbb{Z}[\zeta]$; thus $(a) = \mathfrak{A}^{(1-\jmath)q}$ is also a qth power. Consequently, the abelian extension $\mathbb{K}(a^{1/q})$ is at most ramified at q [Wa, Exercise 9.1, p. 186].

Since $q^2 \mid x$ by (5), we have $a \equiv -\bar{\zeta} \cdot \frac{1-\zeta}{1-\zeta} \equiv 1 \mod q^2\mathbb{Z}[\zeta]$; a fortiori $a \equiv 1 \mod q^2\mathbb{Z}[\zeta, \xi]$. Let $f(X) = \frac{((1-\xi)X+1)^q - a}{(1-\xi)^q}$. Since the binomial coefficients $\binom{q}{i}$ are divisible by q for $i = 1, 2, \ldots, q-1$ and $a \equiv 1 \mod q^2$, the above polynomial is monic with integral coefficients. If β is one of its roots, one easily verifies that it generates the same extension as $a^{1/q}$. The field different divides

$$f'(\beta) = \frac{q}{(1-\xi)^{q-1}} \cdot \big((1-\xi)\cdot\beta + 1\big)^{q-1} \equiv \frac{q}{(1-\xi)^{q-1}} \mod (1-\xi),$$

which is a unit modulo $1-\xi$. It follows that the extension is unramified at q. □

With this, Lemma 9.2 [Wa, p. 109] leads to the proof of Theorem 3:

Proof. Let $\omega \in \bar{\mathbb{Q}}$ be some solution of $X^q = (\alpha/\bar{\alpha})$. By the proof of Theorem 1, the ideal $\mathfrak{A}^{1-\jmath}$ is not principal and since $\mathfrak{A}^{1-\jmath} = ((\alpha/\bar{\alpha})^{1/q}) = (\omega)$, it follows that $\omega \notin \mathbb{Q}(\zeta)$. We let $\mathbb{K} = \mathbb{Q}(\zeta, \xi)$ and $\mathbb{L} = \mathbb{K}[\omega]$ be the unramified abelian extension in the previous lemma; also, $K = \mathrm{Gal}(\mathbb{L}/\mathbb{K})$. Let $\bar{\omega} \in \mathbb{C}$ be the complex conjugate of ω; we show that $\bar{\omega} = 1/\omega \in \mathbb{K}$. Indeed, $(\bar{\omega})^q = (1/\omega)^q$ and thus $\bar{\omega} = \xi^a/\omega$, since the last equation has q solutions in $\mathbb{C}$, differing by qth roots of unity. But $\bar{\omega} \cdot \omega \in \mathbb{R}$ and thus $\xi^a \in \mathbb{R}$, so $a = 0$ and $\bar{\omega} = 1/\omega$. There is thus an element $\jmath \in K$ which acts like complex conjugation. If $\tau \in K$ is the automorphism with $\tau\omega = \xi\omega$, then

$$\tau\jmath\omega = \tau\omega^{-1} = (\tau\omega)^{-1} = \xi^{-1}\omega^{-1} = \jmath(\xi\omega) = \jmath\tau\omega.$$

Thus complex conjugation commutes with K and if L is the group generated by K and $\jmath$, then $\mathbb{Q}[\zeta, \xi]^+ = \mathbb{K}^+$ is the fixed field of $\mathbb{L}$ under L. Let $\mathbb{L}^+ \subset \mathbb{L}$ be the fixed field of $\jmath$; the extension $\mathbb{L}^+/\mathbb{K}^+$ has degree q. If a prime $\mathfrak{p}$ would ramify in this extension, it would have ramification index $q > 2$, which cannot be absorbed by the extension $\mathbb{K}/\mathbb{K}^+$. Therefore $\mathbb{L}^+/\mathbb{K}^+$ is unramified abelian of degree q and by class field theory there is an ideal $\mathfrak{B} \subset \mathbb{K}^+$ which has order q in the class group of $\mathbb{K}^+$ and capitulates in $\mathbb{L}^+$. The theorem follows by symmetry in p and q. □

Acknowledgments

I have learned from Y. Bilu, Y. Bugeaud and G. Hanrot much about the methods which were expanded in this paper. I am grateful to all three of them for enlightening discussions around Catalan's conjecture. I thank R. Schoof for a long and lively email exchange on the local proof and the anonymous referee for his careful reading and helpful suggestions.

References

[Bi] Y. Bilu, Catalan without logarithmic forms (after Y. Bugeaud, G. Hanrot, P. Mihăilescu), submitted for publication, http://www.math.u-bordeaux.fr/~yuri.

[BiH] Y. Bilu, G. Hanrot, Solving superelliptic Diophantine equations by Baker's method, Compos. Math. 112 (1998) 273–312.

[BH] Y. Bugeaud, G. Hanrot, Un nouveau critère pour l'équation de Catalan, Matematika 47 (2000) 15–33.

[BHM] Y. Bugeaud, G. Hanrot, M. Mignotte, Sur l'équation diophantienne $\frac{x^n-1}{x-1} = y^q$, III, Proc. London Math. Soc. 84 (2002) 59–78.

[Ca] J.W.S. Cassels, On the equation $a^x - b^y = 1$, II, Math. Proc. Cambridge Philos. Soc. 56 (1960) 97–103.

[Go] F.Q. Gouvêa, p-Adic Numbers, An Introduction, second ed., Springer, Berlin, 1991.

[Hy] S. Hyyrö, Über das Catalan'sche Problem, Ann. Univ. Turku Ser. A I 79 (1964) 3–10.

[IR] K. Ireland, M. Rosen, A Classical Introduction to Modern Number Theory, second ed., Grad. Texts in Math., vol. 84, Springer, New York, 1990.

[Iw] K. Iwasawa, A note on Jacobi sums, Symp. Math. 15 (1975) 447–459.

[Jh] V. Jha, The Stickelberger Ideal in the Spirit of Kummer with Applications to the First Case of Fermat's Last Theorem, Queen's Papers in Pure and Appl. Math, vol. 93, Queen's Univ. Publ., Kingston, 1993.

[K] C. Ko, On the Diophantine equation $x^2 = y^n + 1$, $xy \neq 0$, Acta Sci. Natur. Univ. Szechuan 2 (1960) 57–64.

[La] S. Lang, Algebraic Number Theory, second ed., Grad. Texts in Math., vol. 110, Springer, New York, 1986.

[Lb] V.A. Lebesgue, Sur l'impossibilité en nombres entiers de l'équation $x^m = y^2 + 1$, Nouv. Ann. Math. 9 (1850) 178–181.

[Mi] M. Mignotte, Catalan's equation just before 2000, in: M. Jutila, T. Metsänkylä (Eds.), Proceedings Inkeri Colloqium, Turku, 1999, de Gruyter, Berlin, 2001, pp. 247–254.

[Mih] P. Mihăilescu, A class number free criterion for Catalan's conjecture, J. Number Theory 99 (2003) 225–231.

[Mih2] P. Mihăilescu, Primary cyclotomic units and a proof of Catalan's conjecture, J. Reine Angew. Math., submitted for publication.

[Pu] J.-C. Puchta, On a criterion for Catalan's conjecture, Ramanujan J. 5 (2001) 405–407.

[Ri] P. Ribenboim, Catalan's Conjecture, Academic Press, San Diego, 1994.

[Wa] L. Washington, Introduction to Cyclotomic Fields, second ed., Grad. Texts in Math., vol. 83, Springer, New York, 1996.